A lesson learnt: the importance of modelling in randomized controlled trials for complex interventions in primary care

Gill Rowlands¹, Jane Sims² and Sally Kerry¹


Background. The Randomised Controlled Trial (RCT) is recognised as the ‘gold standard’ in quantitative research. However RCTs testing health care interventions can be difficult to design and implement. Health care interventions are often complex in themselves and are always applied in complex settings. Such interventions require a process of careful ‘modelling’ to maximize the chances of successful trials that will add to knowledge.

Objectives. To describe the terms ‘complex’ and ‘modelling’ as used in the setting of randomised controlled trials of complex interventions. To give a practical example of an RCT involving a complex intervention applied in a health care setting to illustrate how this might take place in practice.

Methods. We describe an RCT designed and conducted by the authors. We then use our trial as an example to illustrate how complex interventions such as ours might benefit from modelling during the design of the intervention and the setting within which the intervention is to be tested.

Results. Our project was designed and tested before current guidance on complex interventions was published; our RCT was therefore not ‘modelled’ but was based on the outcome of a single quantitative pilot study. As part of our study we ran a parallel qualitative study, which highlighted several areas of complexity both in our intervention, and in the setting within which we applied it. In this paper we show how modelling might have allowed us to recognise these complexities at an early stage and might therefore have resulted in a study more likely to have demonstrated useful outcomes.

Conclusion. Careful modelling of complex interventions is an essential step in designing trials of innovations in health care and health care services. Such a process ensures that interventions fit with and reflect the complexities of the settings within which interventions will be applied, and should ensure that the outcomes chosen are those most appropriate to demonstrate any benefits or risks.

Keywords. Complex interventions, modelling, Randomised Controlled Trial

Introduction

Randomized controlled trials (RCTs) are recognized as the ‘gold standard’ methodology in quantitative research. They can demonstrate the effectiveness or otherwise of interventions in health and practice. Health care interventions are, however, often complex in themselves and are always implemented in complex health care settings.

Thus to maximize the likelihood of adding to knowledge, those planning RCTs must carefully design, model and test their interventions.

In this paper we describe what is meant by the term ‘complex’ in health care research and practice and when and how ‘modelling’ of health care interventions should take place. We illustrate this by describing an RCT we undertook where no objective effect of the intervention was found. The lack of an objective outcome was in contrast to subjective feedback from the study participants who felt that the intervention had produced a change in practice. We describe the process undertaken in designing and running the study and discuss how modelling of the complex processes involved might have led to different study findings.
Importance of modelling in RCTs for complex interventions in primary care

Complexity in health care systems

In their introduction to a series of articles on complexity in health care, Plsek and Greenhalgh define a complex system as one where a collection of individuals have freedom to act in ways that are not totally predictable and whose actions are inter-connected. Such systems are typified by ‘fuzzy boundaries’ where reactions to events reflect factors outside the system and are additionally based on internalized rules and patterns of behaviour. The ‘agents’ within a complex system can change and complex systems themselves are usually influenced by, and have influence on, other systems. Health care systems, with care given by individuals of different backgrounds and professions to individuals with different physical, social and psychological needs, are thus excellent examples of complex systems.

Complexity in health care interventions

Interventions applied in health care can range from those that can be described as simple, such as giving individuals an easily identified single intervention such as a single drug with easily predictable effects, through to very complex multifactorial interventions such as the one described in this paper. The MRC define a complex intervention as one made up of a number of components that may act both independently and inter-dependently (Fig. 1). The authors of the framework describe how it may not be easy to define which elements of a complex intervention are ‘active ingredients’.

‘Modelling’ complex interventions

‘Modelling’ a complex health care intervention prior to testing in the research process enables researchers to gain an accurate understanding of the intervention and its possible effects. The individual components of the intervention need to be delineated and the interaction of these components needs to be assessed and built into the study design. This process will highlight possible outcomes and will thus enable the study team to decide which outcomes will best demonstrate any changes brought about by the intervention. When complex interventions are applied within complex systems, as is seen in health care settings, it is also necessary to learn as much as possible about the interactions between the intervention and the setting within which it will be applied, and to gain an understanding of how that might influence the outcomes of the study. Such modelling can include both quantitative testing, such as computer-based simulations, and qualitative approaches, such as interviews, focus groups or observational studies.

Our study

The study looked at the effect of within-practice educational meetings on GP referral behaviour. The study took place between 1995 (development of hypothesis) through to completion of data collection for the main study RCT in 1999. The study took place in general practices from across London.

The theory underlying our study

The hypothesis was developed following a review of the literature. The key references are highlighted below.

Research into doctors’ learning has shown that doctors’ clinical practice can be improved by educational

---

**Figure 1** Phases in the development of a trial involving a complex intervention. Taken from the MRC framework

<table>
<thead>
<tr>
<th>Theory</th>
<th>Modelling</th>
<th>Exploratory trial</th>
<th>Definitive RCT</th>
<th>Long term implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explore relevant theory to ensure best choice of intervention and hypothesis and to predict major confounders and strategic design issues</td>
<td>Identify the components of the intervention and the underlying mechanisms by which they will influence outcomes to provide evidence that you can predict how they will relate to and interact with each other</td>
<td>Describe the constant and variable components of a replicable intervention AND a feasible protocol for comparing the intervention to an appropriate alternative.</td>
<td>Compare a fully defined intervention to an appropriate alternative using a protocol that is theoretically defensible, reproducible and adequately controlled in a study with appropriate statistical power</td>
<td>Determine whether others can reliably replicate your intervention and results in uncontrolled settings over the long term.</td>
</tr>
</tbody>
</table>
interventions, particularly by those that include peer discussion, rehearsal of communication skills and patient-centred educational activities. In addition, studies have shown that doctors' decisions about whether to perform pathology tests on or write a prescription for a patient can be altered by feedback and by educational meetings. Research into the factors underlying the known wide variability in GP referral rates has highlighted the complexity of the referral decision-making process. Much of that variability is likely to be due to differences in doctor decision-making behaviour. However, when GPs' views were sought on factors influencing the referrals process, they felt that many of these factors might be open to change.

In this study, an intervention was designed to alter GP referral behaviour. It was based on androgogic educational principles (see Box 1) and was applied within general practices. The decision to apply the intervention through educational meetings reflected the design of an intervention that altered doctors' prescribing behaviour. The decision to apply this intervention within general practices reflected the movement in British general practice towards practice-based postgraduate education. The decision to apply this intervention within general practices reflected the movement in British general practice towards practice-based postgraduate education.

The hypothesis we developed following our back-ground reading was that within-practice educational meetings would alter GP referral behaviour and that this would be evidenced by a reduction in GP referral rates.

Development of the study following evaluation of the underlying theory
The development of the study following the 'theory' stage is shown in Figure 2.

Exploratory study. A pilot study was undertaken in one general practice. The pilot study showed an apparent 25% reduction in practice referral rates. This supported the preliminary hypothesis. An RCT was designed and successfully submitted for external peer-reviewed funding.

Definitive RCT. Following the power calculation, 26 practices were recruited to this study and randomly allocated to either an intervention (referrals meetings) or a control (usual referral practice) arm. The referrals meetings are described elsewhere and the guidelines given to practices are shown in Box 2. Referral rates (number of referrals to outside specialists per 100 consultations) were collected per practice for 16 weeks before (T1) and 16 weeks after (T2) the 8 weeks during which the intervention arm practices held their referrals meetings. The change in referral rates (T2–T1) was calculated for each practice. The control and intervention group practices were then compared using an independent samples t-test.

The results of the RCT (Fig. 3) showed that an in-practice educational intervention as applied in this study did not significantly influence practice referral rates. In contrast, feedback from the 62 doctors who participated in the meetings showed a different picture. Sixty doctors returned evaluation forms; of these, 52 (87%) felt that their referral practice had changed, principally through more use of ‘within-practice’ referrals, development of protocols and alteration in factors influencing referrals such as clinical uncertainty and patient pressure for referrals. GPs felt were unwarranted. These findings are described in detail elsewhere.

Parallel qualitative study. In order to examine the processes taking place during the referrals meetings, a qualitative study was conducted in parallel with the RCT. One practice, recruited at the same time as the practices for the RCT, was not entered into the group allocation process. The practice held the referrals meetings in the same way as the intervention arm practices in the RCT but in addition all the meetings were recorded, transcribed and analysed using group dynamic and content analysis.

The qualitative study findings are described in full elsewhere. The analysis highlighted the complexity of the system within which referral decisions were made. First it showed that the partnership was a complex organization made up of individual doctors with differing views on how the practice functioned, especially when decisions were being made, and on what constituted good ‘referral practice’. In addition, the doctors described the impact on their clinical practice of both national and local changes in NHS organization. Each doctor used the opportunity of the protected time within the meetings in different ways. They learnt about their colleagues’ views and feelings about the referral process, discussed the ways in which the referrals decisions had been arrived at, discussed clinical management of difficult cases that had been referred for specialist care and ‘agreed to differ’ about the ways they made their referral decisions.

---

Box 1 Principles of androgogic learning

Androgogic learning is learner centred
Educational needs are identified by the learner
Educational approaches to meet these needs are identified and used by the learner
Any educational approach can be used to meet identified needs but tends to involve approaches such as small group work and role play rather than more traditional approaches such as didactic lectures.
Importance of modelling in RCTs for complex interventions in primary care

Lessons learned: the importance of modelling complex interventions during protocol development

The complex intervention tested in this study was not ‘modelled’ as suggested in the Medical Research Council framework. At the time that this study was designed and conducted, the importance of recognizing and allowing for complexity in health care interventions was only just becoming apparent, leading to the publication of the MRC framework in 2000. This study does, however, provide an excellent illustration of the importance of modelling complex interventions. The findings from doctors’ questionnaire feedback in the RCT combined with the findings from the qualitative arm of the study highlight the complexities in the setting within which the
For the complexities inherent in the intervention and its setting to become clear, and for the assumptions we had made to be identified and challenged, we needed to get GPs’ views and perspectives. This information would have been best gathered through qualitative techniques. One approach would have been to gather information in a stepwise approach with initial questionnaires followed by semi-structured interviews with GPs exploring the questionnaire findings in more depth, and finally a focus group of GPs where interview and questionnaire findings were checked back with GPs. Their views could then have been sought as to the likely impacts of the intervention, how the intervention might vary between practices, and in what ways the impact of the intervention should best be measured.

How might modelling have changed our project? Another action to be taken in modelling the intervention is to delineate the individual components of the intervention, to assess the interaction of these components and to build them into the study design. In our study, this occurred ‘post hoc’ through the qualitative arms of the study and the feedback from the doctors who held the meetings in the RCT (Fig. 4). One way we could have undertaken this process prior to our exploratory study would have been to gather information in a stepwise approach with initial questionnaires followed by semi-structured interviews with GPs exploring the questionnaire findings in more depth, and finally a focus group of GPs where interview and questionnaire findings were checked back with GPs. Their views could then have been sought as to the likely impacts of the intervention, how the intervention might vary between practices, and in what ways the impact of the intervention should best be measured.
In our study, the most important assumptions that might have been altered were that the meetings would have the end result of a reduction in referral rates. Consulting GPs during the design of the study would have told us about what GPs might want to learn during referrals meetings, and what they would regard as appropriate and useful outcomes. In particular, one of our most important findings from the qualitative project—that GPs holding within-practice meetings tended to learn about and accept others’ referral practice rather than challenging referral decisions and learning how to reduce referrals—might have emerged during this exploratory stage. We could then have designed an intervention that would be of maximum benefit for improving referral practice and fashioned outcome measurements that would reflect that. Such outcomes might have been quality of referrals as judged by GPs, specialists or patients, or evidence of completed learning cycles.

**How might modelling have changed our exploratory (pilot) stage?** In the project development process described in the MRC framework document, the exploratory (pilot) stage follows on from and reflects the outcome of the modelling stage of the process. We can build on the hypothesized outcomes of the modelling process described above to picture the sort of exploratory work that might have been undertaken and to compare it with the pilot project we actually conducted.

The pilot project we undertook was not externally funded and was used as the basis of our successful funding bid. Whatever pilot work we conducted would, therefore, of necessity, have been small scale and focused. The findings from the qualitative study and the doctors’ feedback following the RCT indicate that a reduction in GP referral rates would have been an unlikely outcome although we could, of course, still have measured referral rates to assess actual as opposed to perceived impact on referral rates. However, other appropriate outcome measures might have been to look at the use of discussions about GP referrals as an educational tool with completed learning cycles, around learning more either about clinical issues or about colleagues’ referral practice and clinical skills. Alternative outcome measures might have been reviews of the quality of referrals as assessed by key stakeholders such as other GPs, specialists or patients. Doctors’ limited time availability, together with the wealth and mix of skills seen within group practices, would still have made within-practice meetings an attractive proposition. It would have been possible, however, to have run a slightly larger pilot project with one or two referrals meetings in a few practices to explore the meetings and doctors’ views on the likely benefits to arise from a larger study, ways to standardize the intervention whilst maintaining flexibility to maximize benefits and learning opportunities as they took place, and the most appropriate outcome measures to demonstrate any changes seen.

In our project, we moved straight from our hypothesis to a single-practice pilot study. We used referral rates as our preferred outcome measure but did ask for feedback on what the participants had found useful and difficult about the meetings; we used this feedback to decide upon the secondary outcome measures. The reduction in referral rates seen in the pilot project was, in retrospect, almost certainly a chance finding, one that led us into designing a full RCT that was unlikely to demonstrate

---

**TABLE 1**  Modelling our hypothesis: identifying and checking assumptions inherent in the hypothesis

<table>
<thead>
<tr>
<th>Known from previous research</th>
<th>Assumptions in our hypothesis</th>
<th>Unknown factors</th>
<th>How could we have found out more before designing our intervention?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrologic educational interventions improve clinical practice</td>
<td>This applies to GP referrals as well as to other clinical decisions</td>
<td>Are the factors influencing referral decisions the same as for decisions such as prescribing?</td>
<td>GP questionnaires, interviews, focus groups</td>
</tr>
<tr>
<td>Improved practice equates to a reduction in GP prescribing costs and test requests</td>
<td>Fewer referrals equates to better referral practice</td>
<td>GPs’ and specialists views on this. Are there any other measures of referral practice that might be measured?</td>
<td>Individual interviews and focus groups of GPs and specialists</td>
</tr>
<tr>
<td>GPs feel that factors influencing referral practice are open to change</td>
<td>GP’s perceptions on openness to change translates into changed practice</td>
<td></td>
<td>This could only be tested in an RCT</td>
</tr>
<tr>
<td>‘Within-practice’ meetings have the same impact and influence on GPs as meetings held between GPs from different practices</td>
<td>Relationships between GPs within a practice will have similar influences on discussions about referral decisions as relationships between GPs from different practices</td>
<td>Impact of within-practice relationships</td>
<td>Individual interviews of GPs</td>
</tr>
<tr>
<td>External influences (such as the number of consultants) influence GP referral rates</td>
<td>That other external factors will not play a major part</td>
<td>What about relationships with specialists, availability of community services etc?</td>
<td>Individual interviews and focus groups of GPs</td>
</tr>
</tbody>
</table>
the benefits of the intervention we had designed. Conducting the same number of referrals meetings as used in the pilot study (eight) across a few practices instead of just one would have reduced the chance false-positive finding of our pilot and increased the likelihood of identifying measures that would have more accurately reflected the benefits of the intervention.

Discussion

This study provides an example of the importance of following MRC guidance on designing and testing complex interventions in health care settings. Modelling the intervention enables a clearer assessment of any assumptions inherent in the hypothesis and enables the research team to view the intervention in the setting within which it will be applied. This study also, however, demonstrates the difficulty of balancing optimum study design with designing interventions that are practical enough to be applied to everyday practice.

The development of the hypothesis is described above. Given the evidence from previous research into GP clinical behaviour and the organizational changes in the NHS at the time of the development of the study, this preliminary hypothesis was reasonable. In addition, we
undertook a pilot study that appeared to show that our hypothesis and choice of outcome measures were correct and that GPs would find the intervention acceptable. The very low drop-out rate (only one of the 27 participating practices left during the 2 year study) confirms the acceptability of the intervention we designed. We recognized the complexity of the setting and the intervention; we therefore developed secondary outcome measures in the RCT and designed and ran a parallel qualitative arm to the project. In retrospect, however, it is likely that the pilot study we undertook gave a false-positive result leading to a less than optimal intervention design or the choice of an outcome measure insufficiently sensitive to detect changes arising from the study. More sophisticated modelling of the preliminary hypothesis at an earlier stage in the study design process may have uncovered problems and incorrect assumptions inherent in the hypothesis that were later highlighted by the qualitative arm of the study, and might have led to the choice of outcomes measures that would better reflect the benefits of the meetings. It is also possible, however, that such a modelling process may have led to a project design that was more complex with outcome measures that were harder to define and reproduce. The choice of a simple easily measurable outcome (practice referral rates) was obviously attractive both to the funding body and to the practices that participated. If modelling of our hypothesis and intervention had indicated that a different primary outcome measure would be more appropriate, the challenge would then have been to link the changes seen with demonstrable and reproducible ‘hard’ outcomes.

It is also important to consider the possibility that within-practice referrals meetings do not change referral behaviour. We have taken the mismatch between the objective study outcomes and the subjective doctors’ feedback from the meetings as indicating that the meetings did change practice but we simply chose the wrong way to measure that. An alternative possibility is that the meetings indeed had no effect but the doctors involved felt that they had to describe some outcome from a process into which they had invested much time and energy. The research team also developed good relationships with the study practices, again increasing the likelihood of more positive subjective feedback on the meetings than was actually warranted. However, without repeating the study with outcome measures reflecting the participants’ feedback, we have no way of knowing to what extent subjective feedback translated into changed practice.

In conclusion, therefore, careful modelling of complex interventions at an early stage in the development of hypotheses and in the design of projects is an essential stage in developing clinical trials, especially those involving the complexities of human behaviour. We hope that our description of our project, and how it might have been changed and improved by adopting the process described in the MRC framework, will help others as they design and undertake their studies.

Declaration

Funding: The study was funded by the South Thames Region of the NHS Executive.

Ethical approval: The project was submitted to the chair of the Local Ethics Committee who felt it did not require ethics committee approval.

Conflicts of interest: None.

References